

One size fits all? What counts as quality practice in (reflexive) thematic analysis?

Virginia Braun and Victoria Clarke

Qualitative Research in Psychology

Special Issue: Quality in qualitative approaches: Celebrating heterogeneity

Co-edited by Jessica Nina Lester (Indiana University, US) & Michelle O'Reilly (University of Leicester, UK)

Abstract

Developing a universal quality standard for thematic analysis (TA) is complicated by the existence of numerous iterations of TA that differ paradigmatically, philosophically and procedurally. This plurality in TA is often *not* recognised by editors, reviewers or authors, who promote 'coding reliability measures' as universal requirements of quality TA. Focusing particularly on our *reflexive* TA approach, we discuss quality in TA with reference to ten common problems we have identified in published TA research that cites or claims to follow our guidance. Many of the common problems are underpinned by an assumption of homogeneity in TA. We end by outlining guidelines for reviewers and editors – in the form of twenty critical questions – to support them in promoting high(er) standards in TA research, and more deliberative and reflexive engagement with TA as method and practice.

Keywords: Codebook, coding frame, coding reliability, consensus coding, framework analysis, inter-rater reliability, qualitative paradigm, reflexivity, template analysis, theme

Quality in thematic analysis: What matters?

"The authors should discuss how they attempted to avoid bias in their analytic process."

We received this comment in an anonymous review of an empirical paper we had submitted to a specialist qualitative journal, and in which we used our *reflexive* thematic analysis (TA) approach, which we first outlined in this journal in 2006 (Braun & Clarke, 2006). The reviewer's comment reflects a number of problematic assumptions around TA, and indeed qualitative quality criteria, which we contextualise, unpack and expand on in this paper. As the use of TA has expanded, and diversified, since the publication of that paper, the coherence and integrity of published research does not always hold. We use ten problematic practices and assumptions evident in published TA as a tool for thinking about quality in TA, and to support scholars in doing excellent, cohesive TA. These problematic practices reflect confusions and misconceptions about TA that we seek to clarify with clear 'take away' recommendations for quality TA (some of which may apply to qualitative scholarship more broadly, and thus have relevance beyond TA). In order to support editors

and reviewers in their role as quality custodians, we translate these problematic, and corresponding good (or best), practices into 'guidelines' for assessing the quality of TA research – presented in the form of twenty critical questions to consider when evaluating TA research for publication (see Table 1 below).

The problems we commonly encounter in published TA often seem to reflect little to no engagement with the theoretical and philosophical assumptions that underlie procedures, with wider quality discussions in qualitative research (e.g. Levitt et al., 2018; Madill et al., 2000; Sparks & Smith, 2009; Yardley, 2015), or indeed nuanced, aware and situated application of quality criteria. As TA is used within positivist/quantitative *and* qualitative paradigms, and different types of TA are embedded within, and reflect, different conceptualisations of qualitative research, articulating 'universal' quality standards and criteria for TA is challenging. We find Kidder and Fine's (1987) distinction between small q and Big Q qualitative research useful for demarcating between qualitative positivism (small q), the use of qualitative techniques of data collection and analysis within a positivist paradigm, and the use of qualitative techniques within a qualitative paradigm (Big Q qualitative). The values of these are in more or less tension, and without awareness of this tension, published TA research can exemplify what one of our TA co-authors (Nikki Hayfield) memorably dubbed 'confused q' qualitative – research that seems to unknowingly, unreflexively, and incoherently, combine elements of qualitative positivism with the values and assumptions of a qualitative paradigm.

Some might suggest our development of accessible guidelines for 'how to do' (reflexive) TA means we've succumbed to 'proceduralism' (King & Brooks, 2017) or 'methodolatry' (Chamberlain, 2000), that we have *codified* practice, prioritised procedure over theoretical sensitivity and reflexivity, and created rigid and concrete 'rules' for TA research, rather than provided researchers with flexible 'starting points'. This is a misreading of what good (reflexive) TA requires of the researcher. It is the same sort of critique that St. Pierre and Jackson (2014) and others make in relation to 'qualitative coding' – but similarly suffers from a *narrow* reading of what qualitative researching, guided by methods, should and will look like. Although we aim to provide detailed guidelines on process, we also emphasise the fluid, the contextual and contingent, and indeed theory, as crucial within TA. Following procedure is not a guarantor for doing 'good TA'; instead, understanding what the procedures facilitate, what they give you access to, and that these are *tools* for a process, rather than the purpose of analysis, is important. Analytic procedures, including those centred on ensuring and demonstrating quality, typically reflect underlying paradigmatic and epistemological assumptions about meaningful knowledge and knowledge production (Carter & Little, 2007), whether explicated or not. We hope this paper clarifies some potential misuses or misunderstandings. Our aim is to encourage theoretically sensitive (Yardley, 2015), and reflexive and deliberative engagement with TA, of *whatever* variety. By critiquing 'what is,' in terms of TA, we aim to promote 'what could be', and improve both the understanding of the wider terrain of TA, and the enactment of TA in individual research

projects. Although there has been discussion as to whether TA is a distinct method, or a generic set of analytic procedures (e.g., Boyatzis, 1998; Willig 2013), TA does – we believe – offer a distinct way of working with qualitative data, and that, although it shares some features in common with other approaches that seek to identify ‘patterns’ in data (e.g., grounded theory, IPA or qualitative content analysis), it is nonetheless a method (or cluster of methods) in its own right.

(Reflexive) thematic analysis: A brief contextualising introduction

Our aim, when we first wrote about TA (Braun & Clarke 2006), was to articulate an approach to TA that reflected, and was compatible with, the assumptions of a qualitative paradigm (Braun & Clarke, 2019a). This aim has been developed in a qualitative research textbook (Braun & Clarke, 2013), numerous chapters (Braun & Clarke, 2012; Braun et al., 2014a, 2014b, 2016, 2019a; Clarke & Braun, 2016; Clarke et al., 2015; Terry et al., 2017), encyclopaedia entries (Clarke & Braun, 2014a, 2014b), commentaries (Braun & Clarke, 2016, 2019a, 2019b, 2019c; Clarke & Braun, 2018) and editorials about TA (e.g. Braun & Clarke, 2014; Clarke & Braun, 2017). Through this writing, developing a TA website,¹ and teaching, our understanding of the (evolving) landscape of TA (as we see it) has deepened, as has our clarification of where our approach ‘fits,’ and what elements are most vital to quality (reflexive) TA, and why. In recent publications, we have more carefully articulated the assumptions and values around qualitative research that inform our approach to TA (e.g. Braun & Clarke, 2019a) to demarcate what is distinct and different about our approach (e.g. Braun & Clarke, 2019c; Braun et al., 2019a). We now call this approach *reflexive TA* (see Braun & Clarke, 2019a; Braun et al., 2019a, 2019b; Terry et al., 2017). This not only demarcates it as a particular TA approach, it emphasises the importance of researcher’s subjectivity as analytic *resource*, and their reflexive engagement with theory, data and interpretation.

Our original paper sought to provide accessible guidance for TA research that retained flexibility. We emphasised the range of possibilities of different modes of engagement, and our aim was to open-up, rather than close-down, possibilities for TA research, encouraging creativity and wide-ranging use. We partly agree with Potter’s (1997) claim that qualitative analysis is a ‘craft skill’, and something ideally learnt at the feet of ‘master’, something difficult to fully capture through *descriptions* of analytic procedures. However, one of us is a keen crafter, and we know craft skills *can* be formulised for teaching to some extent. Some structured scaffolding can introduce concepts and teach skills to those who do not intuitively just ‘get it’, who are not the proverbial ducks-to-water. In writing about TA, we have sought to provide such scaffolding to learn these skills. This is especially important for democratising access to qualitative methods, as not everyone is fortunate or privileged enough to have access to a ‘master’ qualitative analyst. Interest in qualitative research *far exceeds* the availability of supervisors and remains often under-taught in methods curricula (varying considerably by discipline and by locale). So, we explicated a *six-phase process* for data engagement, coding and theme development. Our most recent articulation of this is: 1)

data familiarisation and writing familiarisation notes; 2) systematic data coding; 3) generating initial themes from coded and collated data; 4) reviewing and refining themes; 5) defining and naming themes; and 6) writing the report. We acknowledge the limits of written guidance, and the potential for it to be (mis)interpreted as prescriptive. However, as noted, there *is* much value in such guidance, not least in making qualitative research more accessible to those without expert supervision or mentoring (McLeod, 2001). Furthermore, we aim to be clear that this phase-approach is *not* intended to be followed rigidly. And as one's analytic (craft) skill develops, these six phases can blend together somewhat, and the analytic process necessarily becomes increasingly recursive.

There are various TA approaches that all aim to identify and make sense of patterns of meaning across a dataset. Despite a shared name and focus on patterned meaning, there are not insignificant differences between different TA approaches (we discuss these later). Our conceptualisation of themes as patterns of shared meaning, cohering around a central concept – the central idea or meaning the theme captures – is not universal, for instance. The *flexibility* of (reflexive) TA as a method, rather than a fully-embedded methodology, means it can be undertaken with quite different guiding theories (albeit constrained by qualitative paradigmatic and epistemological assumptions about meaningful knowledge and knowledge production), and using quite different orientations to data, coding practices and theme development.

Reflexive TA is suited to both experiential (e.g. critical realist, contextualist) and critical (e.g. relativist, constructionist) framings of language, data and meaning (Braun & Clarke, 2013). It can be used for a more *deductive* or more *inductive* analytic process (recognising this can be a continuum, rather than dichotomy). We mean inductive in the sense of analysis 'grounded in' the data, rather than 'pure' induction, because you *cannot* enter a theoretical vacuum when doing TA. Paradigmatic, epistemological and ontological assumptions *inescapably* inform analysis. Researchers using reflexive TA inductively need to identify, and ideally articulate in their reporting, the theoretical assumptions informing their analysis. Using reflexive TA *deductively* means existing research and theory provide the lens through which we analyse and interpret data. Narrowly, this might mean exploring evidence for themes identified in previous research; broadly (and more commonly), this often means using existing political or explanatory theory – such as attachment theory (Willcox et al., 2019) or Foucault's theory of sexual ethics (Beres & Farvid, 2010) – as a lens through which data are coded and interpreted.

Variation also occurs through coding focus, where meaning can be explored across a spectrum from the semantic (surface, obvious, overt) to the latent² (implicit, underlying, 'hidden'). Within reflexive TA, the coding process is integral to theme *development*, in the sense that themes are an 'outcome' of these coding and theme development processes, are developed *through* coding; coding is not – in general – a process for finding evidence for pre-conceptualised themes. The analytic *process* involves immersion in the data, reading, reflecting, questioning, imagining, wondering, writing, retreating, returning. It is far from

mechanical and is a process that requires ‘headspace’ and time for inspiration to strike and insight to develop (Gough & Lyons, 2016). Ho et al. (2017) provide a vivid example of this process of ‘dwelling with’ data, and of “continuously and rigorously reflect[ing] on their own taken for granted thinking” (p. 1760) when researching the experiences of foreign domestic helpers (FDHs) working in Hong Kong, using a hermeneutic phenomenological approach to TA. Ho documents his wondering about unspoken meaning behind the words spoken by participants, reflecting on his personal experiences of hiring FDHs and working as a junior nurse, imagining and questioning how he would feel in situations described and experienced by participants, following insights, looking for instances of similar or contrary language and experiences in the data, and pondering these. It is neither a quick nor an easy process. Time and space (with the data) help develop the nuanced analyses that reflexive TA can deliver, producing rich, complex, non-obvious themes that could never have been anticipated in advance of analysis.

There are some differences between this brief account of (reflexive) TA, and that in our 2006 paper (for discussion of how our thinking has evolved and what has changed, see Braun & Clarke, 2019a; Braun et al., 2019a, 2019b). Our failure to fully articulate the assumptions informing our approach to TA, and how our approach differs from the other approaches we cited (e.g. Boyatzis, 1998), undoubtedly contributes to the confusions and misconceptions apparent in some TA research. We hope *this* paper, alongside other more recent contributions (Braun & Clarke, 2019a; Braun et al., 2019a, 2019b; Terry et al., 2017), serves as a corrective and helps to bring greater clarity, and ultimately assists researchers to avoid the common problems we now outline.

Ten common problems in published TA research

We now highlight ten problems we see in published TA research that cites, or claims to follow, our approach. Such problems are also apparent in TA research more broadly. These problems span broad or conceptual issues, misunderstandings or problematic assumptions, and process or practice problems.

Problem one: Assuming TA is one approach

As previously noted, TA refers not to a singular approach, but rather to a cluster of sometimes conflicting approaches, divergent both in procedure and underlying philosophy, but which share an interest in capturing patterns in data. Yet too often authors do not specify *their* particular orientation to TA, or indeed acknowledge the diversity of TA. We certainly failed to acknowledge this when we first articulated our approach – utilising what was useful, but privately dismissing Boyatzis’ (1998) and other more positivist approaches as ‘not really getting’ the assumptions, values and sensibility of a qualitative paradigm (Braun & Clarke, 2019a). We now think that it is better to conceptualise TA as having several different ‘versions’; we cluster the approaches into what we call “coding reliability”, “codebook” and “reflexive” variations (Braun et al., 2019a). The clustering and demarcation

reflects divergent paradigmatic and epistemological positions and associated procedural differences. Briefly, these are:

- ‘Coding reliability’ TA captures neopositivist approaches that have at their core concerns about ‘objective’ and ‘unbiased’ coding. The use of a codebook for the analytic process, and multiple coders, is key to ensuring ‘accurate’ and ‘reliable’ coding. Such approaches typically use inter-rater reliability (coding agreement) as a key measure of coding quality. They are often deductive in orientation, in the sense that themes are developed early on in, or even prior to, analysis.
- ‘Codebook’ TA captures a cluster of methods that broadly sit within a qualitative paradigm (albeit with some pragmatic compromises). They use some kind of structured coding framework for developing and documenting the analysis, but consensus between coders and interrater reliability are not usually measures of quality. Themes are typically initially developed early on, as they are with coding reliability, but in some methods can be refined or new themes can be developed through inductive data engagement and the analytic process.
- ‘Reflexive’ TA captures approaches that fully embrace qualitative research values and the subjective skills the researcher brings to the process – a research team is not required or even desirable for quality. Analysis, which can be more inductive or more theoretical/deductive, is a situated interpretative reflexive process. Coding is open and organic, with no use of any coding framework. Themes should be the final ‘outcome’ of data coding and iterative theme development.

TA is often written about as if it is just *one* approach. For example, Firmin et al. (2017), in a paper comparing TA and text analysis software, refer to “traditional thematic analysis” (p. 202), as if that is singular and widely understood, but outline a version that is inductive and similar to a modified grounded theory. The assumption of singularity or homogeneity is also at times implicit, with researchers noting their data were analysed using TA, citing multiple (conceptually incompatible or contradictory) approaches, and not providing any specific information about how the analysis was actually conducted – which varies considerably across different versions of TA.

Intimately connected to this ‘one approach’ conceptualisation is a ‘one quality standard’ criterion. This paper’s opening quotation from a reviewer exemplifies this – orienting to quality standards that best apply only to some forms of TA, and in particular to ‘coding reliability’ approaches (e.g. Boyatzis, 1998; Guest et al., 2012; Joffe, 2012). These approaches have the longest history, and are often similar to TA’s close cousin, qualitative content analysis (e.g. Forman & Damschroder, 2008). Boyatzis (1988) offered his approach to TA as one to ‘bridge the divide’ between positivist (quantitative) and interpretative (qualitative) paradigms by marrying the collection and analysis of qualitative data with positivist markers of quality – reliability of measurement or observation and containing researcher ‘bias’ through the standardisation of coding procedures and the demonstration of consensus among multiple coders. Such quality markers are often assumed to apply to *all*

forms of TA. And yet the logic, process, and thus quality measures for coding reliability TA are quite different from reflexive TA. It is important that qualitative research, and different approaches to qualitative research, are evaluated on their own terms (Madill et al., 2000; Sparks & Smith, 2009; Yardley, 2015).

As briefly noted above, these coding reliability approaches differ from reflexive TA in recommending *early* theme development (perhaps following some data familiarisation), the use of a structured and fixed codebook or coding frame (perhaps developed following some data familiarisation or initial analysis of a portion of the data), the use of multiple coders who work independently, measurement of between-coder agreement (or inter-rater reliability) and the determination of final coding through consensus. The over-riding concern is with demonstrating the accuracy or reliability of coding, and this positivist prioritisation shapes how analysis is conducted. Demonstrating coding reliability and the avoidance of 'bias' is illogical, incoherent and ultimately meaningless in a qualitative paradigm and in reflexive TA, because meaning and knowledge are understood as situated and contextual, and researcher subjectivity is conceptualised as a *resource* for knowledge production, which inevitably sculpts the knowledge produced, rather than a must-be-contained threat to credibility. The application of such 'coding reliability' criteria to reflexive TA also, to us, suggests that the researcher does not fully 'get' the *fundamentals* of reflexive TA, does not understand what the qualitative values underpinning the framework expect or delimit.

Between these two 'poles' of TA, 'codebook' approaches, like framework analysis (Gale et al., 2013; Ritchie & Lewis, 2003; Smith & Firth, 2012), matrix analysis (Miles & Huberman, 1994; Nadin & Cassell, 2014) and template analysis (e.g. Brooks, McCluskey, Turley & King, 2015; King, 2012, 2014), comprise processes and conceptualisations that have elements of each, with their own 'best practice' and quality criteria guidelines. They share early theme development (of some or all themes) and the use of a structured codebook or coding frame (the framework, template or matrix) with coding reliability approaches, and the qualitative philosophy and values of reflexive thematic analysis, such as recognising researcher subjectivity and that knowledge is contextual (see Braun et al., 2019a, for more discussion). For some codebook proponents, their approach represents some degree of 'compromise' of qualitative principles, with research driven by pragmatic demands around pre-determined information needs (with 'themes' often consisting of summaries of responses to particular questions), strict time frames for producing 'results,' and the necessity of team work. Multiple researchers code different portions of the data, facilitating delivery of 'results' to a fixed deadline. The methods facilitate inclusive teams of researchers, opening participation for qualitative novices and participants or stakeholders, with little or no research background. Data are often rather concrete, and the required output can often be a descriptive or summative analysis of semantic meaning with results accessible to and actionable by practitioners and stakeholders (Ritchie & Spencer, 1994; Smith & Firth, 2011). These approaches often demonstrate a 'qualitative pragmatism' and work well for applied research – for instance, the framework approach was developed by researchers at the

(British) National Centre for Social Research in the 1980s for use in applied social policy research (Ritchie & Spencer, 1994).

The take away... 'TA' is best thought of as an umbrella term for, or a 'fuzzy' set (Madill & Gough, 2008) of, approaches that share some characteristics in common (analysis through coding and theme development; some degree of theoretical and research design flexibility; a focus on semantic and latent meaning) but can differ significantly in both underlying paradigmatic and epistemological values, and in procedures. It is vital that researchers, reviewers and editors understand the broad paradigm distinctions between different versions of TA. We encourage TA researchers to clearly demarcate *which* TA approach they are using. Furthermore, if they cite authors from different orientations to TA, to clearly specify what they are 'taking' from each and justify (well) any use of divergent criteria and practice.

Problem two: Citing without reading!

Unfortunately, this problem is not a joke. Some citations of our 2006 paper appear performative: dropped in as a 'convenient' or maybe even 'required' citation. Numerous authors claim to have 'followed the procedures outlined by Braun and Clarke (2006),' then describe procedures with little or no resemblance to those we outlined. For example, Kaye, Wall and Malone (2016) described their approach as TA and their procedure as "in line with Braun and Clarke (2006) analytic strategy" (p. 464), but then outlined an analytic process more akin to coding reliability TA: the data were read by "two naïve coders"; each coder independently identified initial themes; the data were coded to test the validity of the initial themes (in part to determine the frequency of each theme); the data were reviewed again alongside the development of a codebook. Reading such papers, we have discovered that we promote the use of codebooks and coding frames, consensus coding, the measurement of coding reliability, developing themes before data coding, data or theme saturation, the measurement and reporting of theme frequency, constant comparative analysis, and more... Reader, we do not! Not only are these things we have *not* said, they are all things we are indeed critical of, as practices for Big Q qualitative inquiry (Braun & Clarke, 2013, 2019c; Clarke & Braun, 2019). The most plausible (and perhaps generous) explanation for claims that we advocate for procedures that we do not in fact advocate for, is that the authors have not read our paper.

The take away... Do your homework – by which we mean, do not cite us (or indeed any other TA author) as the method used for TA without: a) reading the methodological guidelines provided; b) confirming that what you did is what is advocated; and c) being clear on why any 'deviations' from the broad approach were adopted.

Problem three: Unjustified or incompatible 'mash-ups'

There are numerous instances of problematic practice that go beyond researchers citing multiple (incompatible) sources for TA, effectively combining reflexive TA with other often incompatible procedures without justification or explanation. The notion of methodological

'mash-ups' is not problematic per se; the flexibility of TA invites such creativity and innovation. Our concern is with seemingly unknowing and unreflexive 'mash-ups' that result in theoretical and conceptual incoherence – 'confused q' research. For example, the combination of reflexive TA with the use of codebooks and coding reliability measures is common, but the *tensions* between the organic and subjective coding processes of reflexive TA and concerns for coding accuracy and reliability seem to be far less commonly recognised – certainly they are not often discussed. Some researchers also use concepts and terminology distinctively associated with other analytic approaches (e.g. the terms emergent and superordinate theme, associated with interpretative phenomenological analysis [IPA]; the grounded theory concepts of constant comparative analysis, line by line coding and saturation; see Braun & Clarke, 2019c, for a critical discussion of the use of the saturation concept in TA) without discussion of the located and particular meanings, or theoretical anchors, associated with these concepts and practices. They do not always translate (well) to, or cohere with, TA. We also see researchers 'supplementing' (reflexive) TA with additional procedures or approaches, arguing that TA is insufficient in and of itself to achieve their analytic purpose (e.g. Floersch et al., 2010) – which seems to be based on flawed assumptions and a singular idea of what TA is (discussed further in subsequent problems).

The take away... Methodological 'mash-ups' should be warranted, justified and theoretically coherent, and based in a full understanding of what (reflexive) TA can – and cannot – offer.

Problem four: Assuming TA is atheoretical

The theoretical flexibility of TA, the absence of *inbuilt* guiding theory, is where TA departs from other popular and well-utilised qualitative analytic approaches such as IPA, grounded theory, discourse analysis and narrative analysis. Some of these approaches offer a range of theoretical possibilities through the proliferation of different iterations of the original approach (e.g. there are various 'flavours' of grounded theory – positivist, contextualist/constructivist and radical constructionist are all discussed; see Charmaz & Henwood, 2008; Glaser, 1992; Pidgeon & Henwood, 1997). The relative lack of theoretical prescription inherent in TA is often misinterpreted as indicating TA is atheoretical. We also encounter TA being *treated as* an atheoretical method through researchers failing to specify the theoretical assumptions informing their engagement with TA. Indeed, as became clear to us in an author's response to our review of their manuscript recently, it is sometimes assumed that inductive TA is entirely without theoretical foundations, and that only *deductive* TA requires discussion of theory. (We also see *deductive* TA misconceptualised as TA informed by a research question, or the researcher's ontological and epistemological assumptions or interests in the topic.)

Despite not having *inbuilt* theory, TA can *never* be conducted in a theoretical vacuum; researchers always make assumptions about what data represent (e.g. do participants' words relatively transparently communicate individual experience or do their words constitute social discourse, performing social actions?), what can be claimed on the basis of

these data, and indeed what constitutes meaningful knowledge. If the assumptions made are more ‘common-sensical’ or reflect the dominant assumptions within a discipline, it can be hard to recognise these *as* assumptions, indeed *as theory*, but they are nonetheless theoretically-informed assumptions with consequent analytic implications.

The take away... Researchers should always reflect on and specify the philosophical and theoretical assumptions informing their use of TA, even inductive TA. TA should be recognised as a (more or less, depending on the specific iteration) theoretically flexible, but not atheoretical, approach and one *equally* suited to experiential and critical orientations for qualitative research (as we now discuss, TA is often assumed to be only appropriate to use in experiential orientations).

Problem five: Assuming TA is only realist/essentialist or experiential/phenomenological

Closely connected to problem four, TA is regularly positioned implicitly as a realist or essentialist method that simply retrieves truth and reality, both of which are treated as unproblematically accessible (and are often un[der]theorised). Experience is likewise often assumed to be accessible through TA, and TA is commonly described as *particularly* compatible with phenomenology (e.g. Guest et al., 2012; Joffe, 2012) or even as a phenomenological method: “thematic analysis adopts a phenomenological position to systematically identify themes” (Newton-John et al., 2017: 1822). Before IPA, TA was used as a phenomenological method in psychology (e.g. Dapkus, 1985), yet the proclamation that TA and phenomenology are aligned is rarely explained. We speculate that this reflects an understanding of TA as (only) compatible with broadly experiential approaches to qualitative research, and the analysis of “subjective viewpoints” (Flick, 2014: 423) – research underpinned by a reflective view of language and focused on exploring participants’ lived experience, sense-making, views, needs, practices and so on, through a broadly ‘empathic’ lens (Braun & Clarke, 2013; Willig, 2013). This framing is unnecessarily limited.

In contrast, we position (reflexive) TA as an approach that is flexible enough to be compatible with both experiential *and* critical qualitative research (see Clarke & Braun, 2014b) – no doubt a reflection of our background and training in critical qualitative psychology (see Braun & Clarke, 2019a; Lainson et al., 2019; Jankowski et al., 2017). Moreover, there is a tradition of ‘thematic’ discursive methods in psychology such as thematic decomposition (Stenner, 1993) and thematic discourse analysis (Singer & Hunter, 1999; Taylor & Ussher, 2001). And we increasingly see exciting ‘mash-ups’ of reflexive TA and discursive and narrative approaches, and the development of hybrid approaches like ‘thematic narrative analysis’, which combines TA to identify themes with narrative analysis to explore the sequential organisation of events in participants’ accounts (e.g. Palomäki, Laakasuo & Salmela, 2013). ‘Critical thematic analysis’ (e.g. Terry & Braun, 2011) likewise combines reflexive TA with some features of critical discursive psychology (e.g. Wetherell & Edley, 2008). It is precisely this kind of theoretically-knowing, creative and reflexive ‘mash-ups’ that we welcome, in contrast to the seemingly *unknowing* ‘mash-ups’ of reflexive TA with qualitative positivism described in problem three.

The take away... Avoid treating TA as if it inherently offers only *one type of* orientation to qualitative research. Instead, provide a rationale that explains the particular use of (reflexive) TA, and the particular orientation to (reflexive) TA you are taking.

Problem six: Assuming TA is only descriptive

Closely related to the previous two problems is the notion that TA is only a descriptive or data reduction method, in which data patterns are paraphrased or summarised. TA research is assumed to offer a low level of interpretation compared to approaches such as grounded theory or narrative analysis (e.g. Aguinaldo, 2012; Vaismoradi et al., 2013), and this impoverished conceptualisation of TA is often used to justify combining TA with other approaches (such as grounded theory) that are positioned *as* offering (a higher level of) interpretation (e.g. Floresch et al., 2010). We have two concerns with this problematic assumption.

First, description and interpretation are positioned as separate and distinct activities. And in descriptive or summative analyses, the researcher appears to become a passive, disinterested and decontextualised conduit for the voices of participants. We contend that even TA with a descriptive purpose is an *interpretative* activity undertaken by a researcher who is situated in various ways, and who reads data through the lenses of their particular social, cultural, historical, disciplinary, political and ideological positionings. They edit and evoke participant “voices” but ultimately tell *their* story about the data: “social research cast through voices typically involves carving out unacknowledged pieces of narrative evidence that we select, edit, and deploy to border our arguments” (Fine, 1992: 218). And positivist-empiricist reporting practices, like the passive voice (of the objective scientist), and indeed appeals to demonstrating ‘coding reliability’ – for ‘accuracy’ – can obfuscate *our responsibility* for the ways we interpret participants’ accounts (Clarke & Braun, 2019). Our language use is never neutral, even in apparently descriptive reporting.

Second, the conceptual reduction of TA to a low-level descriptive method obscures the range of possibilities reflexive TA holds – most particularly its potential for deeply interpretative, theorised analyses. If such possibilities and potentials were recognised and fully appreciated, researchers could stop adding another (supposedly more interpretative) method to ‘extend’ their TA – such ‘mashups’ are often not just unnecessary, but unconvincing or even problematic.

The take away... Interpretation is inherent to the (TA) analytic process, and there is nothing in the method of TA that renders it simply summative or descriptive. Interpretative depth lies in the skill of the analyst, not the method. Researchers supervising (post)graduate student research can usefully emphasise this point to students. (We get many emails from anxious students who have been told that TA is unsuitable for postgraduate and especially doctoral research because it lacks sophistication and is *only* an atheoretical, naïve realist, descriptive method.)

Problem seven: Confusing codes and themes

Reflexive TA makes a distinction between codes and themes, but there is no *absolute* distinction between codes and themes across TA research. In many TA approaches, these terms are used interchangeably, or coding is conceptualised as a *process* of allocating data to predetermined themes. In reflexive TA, a code is conceptualised as an analytic unit or tool, used by researcher to develop (initial) themes. Here, codes can be thought of as entities that capture (at least) one observation, display (usually just) one facet; themes, in contrast, are like multi-faceted crystals – they capture multiple observations or facets (occasionally, rich, complex and multifaceted codes might be ‘promoted’ to themes [Charmaz, 2006], a process called ‘subsumption’ in IPA [Smith et al., 2009]). In TA papers where numerous ‘themes’ are presented, the ‘themes’ are often one dimensional and meaning-thin; they tend to capture only one (small) observation or facet of meaning (and quite often they are named with just one word). For example, Fornells-Ambrojo et al. (2017), in mixed methods research on service users’ experiences of routine outcome monitoring (ROM) in an improving access to psychological therapies for severe mental illness service, reported seven ‘overarching’/‘superordinate’ themes (‘superordinate’ is associated with IPA [Smith et al., 2009] not TA) and 18 themes. The overarching themes were nested under two headings: 1) helpful and 2) unhelpful aspects of ROM. From the information presented, the themes appeared to each capture a single semantic observation or insight about users’ perceptions (e.g. expressing my feelings, being understood). These are better conceptualised as (reflexive TA) *codes*. The overarching or superordinate themes were similarly ‘thin’, and mostly named with one word (e.g. format, distressing, disliked). This also illustrates ‘analytic foreclosure’ (Connelly & Peltzer, 2016) – with further analytic work, underlying patterns of shared meaning could have been identified that drew together several of these ‘themes’ (codes) into richer, more complex themes that revealed multiple facets of a particular meaning or experience. (As a pre-defined purpose of the analysis was to identify positive and negative aspects of ROM, a codebook or coding reliability approach to TA might have been more in keeping with the purpose of the study, depending on paradigmatic positionings, which in the study leaned towards the positivist, as inter-rater reliability was measured.)

The take away... Is it a code, or is it a theme? Clarification of what codes and themes represent, and what role they play in analysis, is important. Researchers should consider whether their provisional or candidate themes are one- or multi-dimensional and ensure their conceptual frameworks and reporting of analytic outputs align with the version of TA they are using.

Problem eight: Confusing themes and topics

This is probably one of the most common problems in published TA research, and it hinges on confusion around how themes are conceptualised, as there is no one widely agreed on and adopted conceptualisation of a theme in TA (DeSantis & Ugarriza, 2000). We have subsequently clarified our conceptualisation of themes, because it was evident that our initial definition of a theme as capturing ‘some level of patterned response or meaning’

(Braun & Clarke, 2006) left too much room for confusion. Themes in reflexive TA are patterns of shared meaning, united by a central concept or idea (Braun & Clarke, 2013, Braun et al., 2014). This means themes might draw together data that on the surface appear rather disparate. As previously noted, themes are also multifaceted. We like to think of themes as stories – stories we tell about our data. Data topics (sometimes called domains) are not themes in this way – they are things discussed in (say) an interview, perhaps introduced in a question from the interview guide. Participants quite often provide divergent and disparate responses around a topic, but it is common for summaries of topics or domains to be presented as *themes*. In these data-topics-as-themes, the participants responses are summarised, but there is no central concept, no *shared* meaning, only a shared *topic*.

To give an example, Senders et al. (2014), in research exploring patient perceptions of how stress is addressed in medical treatment for multiple sclerosis, which used TA “according to Braun & Clarke” (p. 1678), but also involved the use of codebooks and consensus coding, reported two themes: 1) facilitators and 2) barriers to talking about stress in the medical visit. Each theme incorporated the same three sub-themes: the medical system, clinician behaviour and patient behaviour, that each facilitated or inhibited communication about stress. What is reported is a series of observations related to the topics of, for instance, facilitators or barriers and the medical system. This type of analysis seems better suited to a codebook or coding reliability TA, in which themes can be predetermined prior to analysis and themes *may* consist of summaries of data domains. To provide a clear example of shared-meaning themes, Tischner (2019), in her story completion study exploring constructions of weight loss motivations and health, presented five themes, including one entitled ‘weight-loss activity as good for every woman’. The title alone clearly conveys that this is a shared meaning with a central organising concept theme – this highlights the importance of naming themes well (Braun & Clarke, 2013). Tischner’s theme captured the way weight loss was positively framed as a way of improving appearance for women and occupied a mostly unquestioned position as desired and desirable by and for every woman, and something all women would inevitably engage in at some point in their lives.

What is no doubt confusing is that some TA approaches, particularly coding reliability and codebook approaches, treat such topic summaries *as* themes, when these are quite different from understandings of patterns of thematic-meaning. Indeed, if themes are developed prior to any analysis and coding, it is unlikely they can be much more than topic summaries, as it’s difficult to conceive of the type of thematic patterning that is the outcome of coding, built from codes, being fully anticipated in advance of any or much analytic work taking place. For themes to be patterns of shared meaning underpinned by a central concept, they must be analytic outputs, not inputs. In addition to the conceptual confusion around ‘what is a theme’, the use of topics as themes can also be another example of analytic foreclosure (Connelly & Peltzer, 2016).

The take away... We encourage researchers to reflect on their understanding of themes, to use themes-as-shared-meaning and themes-as-shared-topic knowingly and reflexively, and to clearly justify *any use of* topic summaries for 'themes' in reflexive TA.

Problem nine: Emerging themes – confusing 'themes-as-pre-existing analysis' with 'themes-as-the-outcome of analysis'

The way some researchers and methodologists write about themes, they appear to conceptualise them as entities that pre-exist analysis, lurking about in the dataset. The researcher's task is to locate and retrieve these themes, for reporting. Such themes are "diamonds scattered in the sand" (Braun & Clarke, 2016: 740), found by the researcher, rather than actively created by them through their interpretative engagement with data. This is effectively "thematic discovery," as Guest, Bunce and Johnson (2006: 66) dubbed it. This notion of themes as diamonds in the sand is evident in the phrase 'themes emerged', used by countless authors of TA papers, and evoked by critics of qualitative methods like TA, as *the process* of theme development (e.g. "thematic analysis in which themes somehow miraculously emerge from the data"; St. Pierre, 2019: 4). We appreciate that the concept of 'emergent themes' is used in approaches such as IPA (and sometimes TA) to reflect the inductive creation of themes, but we are troubled by the implications of the claim 'themes emerged'. The phrasing evokes a process that *suggests* that themes present from data with little intervention from the researcher other than extraction, once the themes reveal themselves from the (potentially murky) data depths.

We were critical of the language of 'themes emerge' in our 2006 paper: "An account of themes 'emerging' or being 'discovered' is a passive account of the process of analysis, and it denies the active role the researcher always plays in identifying patterns/themes, selecting which are of interest, and reporting them to the readers" (p. 80). We quoted Ely et al. (1987: 205-6) who argued that "If themes 'reside' anywhere, they reside in our heads from our thinking about our data and creating links as we understand them." It seems that our argument was not particularly persuasive, as so many researchers citing our paper and claiming to follow our approach, refer to 'the themes that emerged' from their data (this connects, no doubt, to problem two). We acknowledge that our phrasing of the third phase of reflexive TA – 'searching for themes' – has likely contributed to confusion around the conceptualisation of themes as pre-existing entities that reside in data. For this reason, we have, for now, relabelled this phase 'generating initial themes' to highlight the active role of the researcher in theme creation *and* the provisionality of themes when first developed. A 'pre-existing analysis' conceptualisation of themes can *perhaps* be reconciled with conceptions of topics as themes and themes as analytic inputs, and some TA proponents do appear to conceptualise themes as entities that pre-exist analysis, but this does *not* align with reflexive TA.

The take away... We encourage researchers using reflexive TA to write about theme generation as a creative and active process, one they are central to, and to *always* avoid claiming that themes emerged.³

Problem ten: Uncritical acceptance of what we say

The final problem we want to highlight could, indeed, be called proceduralism. But we hope our implicit and explicit message to ‘be a critical, thinking researcher and writer’ makes this point not seem contrary. Our emphasis on quality involves discouraging slavish or unthinking adherence to procedures, and any accusation of that represents a misconceptualisation of our articulation of TA (and qualitative research more widely). Rather, we encourage theoretically-knowing, reflexive and ‘aware’ use of TA; we task researchers with appreciating the diversity and flexibility of TA, and the ways analytic and quality procedures reflect paradigmatic and epistemological assumptions. We offer qualitative researchers reflexive TA as a flexible ‘starting point’ for theoretically sensitive and creative research and invite them to make it their own. To do good reflexive TA, choices related to theory, data orientation, and more, *must* be made, articulated and coherently enacted.

The take away... Be a thoughtful researcher; do not just slavishly follow what methodology writers say. We do not provide a full holiday package; we provide a compass and a map to navigate your adventure (Braun et al., 2019b). And we encourage researchers to reflect on whether it is a reflexive, coding reliability or codebook version of TA that is best suited to their philosophical commitments and the analytic purpose of their research.

Introducing our evaluation tool for reviewers and editors

Researchers who do TA have an important responsibility to do the best quality research they can. But published work also goes through quality assurance processes – peer review, and editorial decision making and guidance – and these steps are also important for ensuring quality. Unfortunately, we know from personal experience they can also be the points at which problems (including some of the 10 discussed above) can either creep in, or even be demanded. Authors can be in a tricky spot navigating the review process, and if reviewers or editors suggest or demand something problematic, authors have to: 1) recognise that it is problematic; and 2) find a way to navigate through it. What is the author of a reflexive TA study to do, when faced with a reviewer or editor comment like the one we started with? Should they attempt to ‘educate’ reviewers and editors about the plurality of TA, the assumptions underpinning reflexive TA (Levitt et al., 2018)? Should they highlight the lack of a single quality standard that cuts across all qualitative analytic approaches? Yes! We encourage authors to *explain* and *defend* their research values, using the information provided in this paper and elsewhere to justify their challenges to requests or requirements from editors and reviewers.

However, authors can only argue so far... And editors and reviewers hold ultimate responsibility for enacting (or failing to enact) the quality standards debated and developed in the wider research community. There are some basic editorial practices that can facilitate the publication of good quality TA: editors should understand not just their own methodological limits, but their theoretical/conceptual and methodological assumptions

and values; editors should select at least one reviewer with appropriate methodological expertise to review TA manuscripts (Levitt et al., 2018); and editors should ensure the journal's editorial board includes at least one qualitative methodologist with wide-ranging expertise who can help guide appropriate reviewer selection – this latter point is vital if the editor is not a qualitative methodologist. Unless at least one of the 'expert' reviewers chosen by the editor has very wide-ranging and in-depth knowledge of TA, it is easy to see how *poor* practice is not questioned or challenged, and how the problems we have documented have become so *common*.

As authors, we can not only request, but also encourage, this. We can embed our responses to reviews and editorial feedback in 'best practice' literature. To facilitate better editorial (and reviewer) practice in relation to TA, we have developed a list of evaluation questions for editors and reviewers that lay out some of the key tensions and best practices related to TA (see Table 1). These critical questions can offer a quick and easy resolution to at least some of published problems that we see, and we encourage TA authors to bring these to editors' attention.

[Insert Table 1 about here]

Finally, we recommend that journal editors provide *longer* page limits for TA – and other qualitative – papers (Levitt et al., 2017, 2018). Short word counts can significantly constrain how TA research is reported, which can contribute to apparently poorer practice, in all sorts of ways. As Levitt et al. (2018) argued, quality expectations and contextualisation, research reflexivity and illustrating findings with data extracts *demand* more manuscript pages. As journals are now primarily digital, the scope for expansion is increased. But we would not advocate for a solution where data or methodological commentary are relegated to secondary tables or additional online appendices, because we see these as *integral* to quality judgements and process (Braun & Clarke, 2019b).

Conclusion

In order to discuss quality in TA, we have delineated ten common problems in published TA research that cites, or claims to follow, our reflexive approach to TA. The first of these – assuming TA is one approach – underlies most of the other problems. And indeed, it is the diversity and plurality of TA – that TA ranges from positivism to critical qualitative paradigms – that presents a key challenge for the qualitative research community in demarcating quality standards for TA research, an issue that has dogged wider qualitative quality discussions. To improve the quality of published TA, we encourage researchers to reflect on the relationship between analytic practices, including quality practices, and the ontological and epistemological foundations of research, and to use TA knowingly, deliberately and reflexively. We task reviewers and editors – who are effective arbiters of research quality – with supporting researchers in realising this. To this end, we have provided twenty critical questions to consider when reviewing or editing TA manuscripts. But like everything we write about TA, this reflects our *current* thinking, and things change. So, our most vital piece

of advice for anyone using TA, whether reflexive or another approach, is to read the most up to date writing and advice from authors, rather than just the 'key reference' for an approach. And finally, we emphasise that TA is not a method for all purposes! Instead of trying to make it fit, when it does not, explore what the many other wonderful qualitative analytic approaches might offer your project instead.

Notes

¹ <https://www.psych.auckland.ac.nz/en/about/our-research/research-groups/thematic-analysis.html>

² Latent should not be understood as – only – referring to unconscious meaning, a common misconception among relationally and psychoanalytically oriented psychological practitioners.

³ We have chanted the mantra 'themes don't emerge' so often the Etsy shop Science on a Postcard has created a 'Themes. Do Not. Emerge.' badge. And we're happy to see it's a 'bestseller'! <https://www.etsy.com/uk/listing/635756651/themes-do-not-emerge-qualitative>.

References

- Aguinaldo, J.P. (2012). Qualitative analysis in gay men's health research: Comparing thematic, critical discourse, and conversation analysis. *Journal of Homosexuality, 59*(6), 765-787.
- Beres, M. A. & Farvid, P. (2010). Sexual ethics and young women's accounts of heterosexual casual sex. *Sexualities, 13*(3), 377-393.
- Boyatzis, R. E. (1998). *Transforming qualitative information: Thematic analysis and code development*. Thousand Oaks, CA: Sage.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*, 77.
- Braun, V., & Clarke, V. (2012). Thematic analysis. In H. Cooper (Ed.), *APA Handbook of Research Methods in Psychology* (Vol. 2: Research Designs, pp. 57-71). Washington, DC: APA books.
- Braun, V., & Clarke, V. (2013). *Successful qualitative research: A practical guide for beginners*. London: Sage.
- Braun, V. & Clarke, V. (2014). What can 'thematic analysis' offer health and wellbeing researchers? *International Journal of Qualitative Studies on Health and Well-being, 9*, 10.3402/qhw.v9.26152
- Braun, V. & Clarke, V. (2016). (Mis)conceptualising themes, thematic analysis, and other problems with Fugard and Potts' (2015) sample-size tool for thematic analysis. *International Journal of Social Research Methodology, 19*(6), 739-743.
- Braun, V. & Clarke, V. (2019a). Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise & Health, 11*(4), 589-597.
- Braun, V. & Clarke, V. (2019b). Novel insights into patients' life-worlds: The value of qualitative research. *Lancet Psychiatry, 6*, 720-721.
- Braun & Clarke, V. (2019c). To saturate or not to saturate? Questioning data saturation as a useful concept for thematic analysis and sample-size rationales. *Qualitative Research in Sport, Exercise & Health*.
- Braun, V., Clarke, V. & Hayfield, N. (2019b). "A starting point for your journey, not a map": Nikki Hayfield in conversation with Virginia Braun and Victoria Clarke about thematic analysis. *Qualitative Research in Psychology*.
- Braun, V., Clarke, V. & Rance, N. (2014). How to use thematic analysis with interview data. In Vossler, A. & Moller, N. (Eds.), *The Counselling & Psychotherapy Research Handbook* (pp. 183-197). London: Sage.

- Braun, V., Clarke, V., & Terry, G. (2015). Thematic analysis. In P. Rohleder & A. Lyons (Eds.), *Qualitative Research in Clinical and Health Psychology* (pp. 95-113). Basingstoke: Palgrave MacMillan.
- Braun, V., Clarke, V. & Weate, P. (2016). Using thematic analysis in sport and exercise research. In B. Smith & A. Sparkes (Eds.), *International handbook on qualitative research in sport and exercise*. London: Routledge.
- Braun, V., Clarke, V., Terry, G & Hayfield N. (2019a). Thematic analysis. In Liamputtong, P. (Ed.), *Handbook of research methods in health and social sciences* (pp. 843-860). Singapore: Springer.
- Brewster, M.E., Velez, B.L., Mennicke, A. & Tebbe, E. (2014). Voices from beyond: A thematic content analysis of transgender employees' workplace experiences. *Psychology of Sexual Orientation and Gender Diversity, 1*(2), 159-169.
- Brooks, J., McCluskey, S., Turley, E. & King, N. (2015). The utility of template analysis in qualitative psychological research. *Qualitative Research in Psychology, 12*(2), 202-222.
- Carter, S.M. & Little, M. (2007). Justifying knowledge, justifying method, taking action: Epistemologies, methodologies, and methods in qualitative research. *Qualitative Health Research, 17*(10), 1316-1328.
- Chamberlain, K. (2000). Methodolatry and qualitative health research. *Journal of Health Psychology, 5*(3), 285-296.
- Charmaz, K. (2006). *Constructing grounded theory: A practical guide through qualitative analysis*. Thousand Oaks, CA: Sage.
- Charmaz, K., & Henwood, K. (2008). Grounded theory. In C. Willig & W. Stainton Rogers (Eds.), *The Sage handbook of qualitative research in psychology* (pp. 240-259). London: Sage.
- Clarke, V. & Braun, V. (2014a). Thematic analysis. In A. C. Michalos (Ed.), *Encyclopaedia of quality of life and well-being research* (pp. 6626-6628). Springer, Dordrecht, Netherlands: Springer.
- Clarke, V. & Braun, V. (2014b). Thematic analysis. In T. Teo (Ed.), *Encyclopaedia of critical psychology* (pp. 1947-1952). New York: Springer.
- Clarke, V. & Braun, V. (2016). Thematic analysis. In E. Lyons & A. Coyle, (Eds.), *Analysing Qualitative Data in Psychology, 2nd ed.* (pp. 84-103). London: Sage.
- Clarke, V. & Braun, V. (2017). Thematic analysis. *Journal of Positive Psychology, 12*(3), 297-298.
- Clarke, V. & Braun, V. (2018). Using thematic analysis in counselling and psychotherapy research: A critical reflection. *Counselling and Psychotherapy Research Journal, 18*(2), 107-110.
- Clarke, V. & Braun, V. (2019). Feminist qualitative methods and methodologies in

psychology: A review and reflection. *Psychology of Women and Equalities Section Review*, 2(1), 13-28.

Clarke, V., Braun, V., & Hayfield, N. (2015). Thematic analysis. In J. Smith (Ed.), *Qualitative psychology: A practical guide to research methods (3rd ed.)* (pp. 222-248). London: Sage.

Connelly, L. M. & Peltzer, J. N. (2016). Underdeveloped themes in qualitative research: Relationships with interviews and analysis. *Clinical Nurse Specialist*, January/February, 51-57.

Crabtree, B., & Miller, W. (1999). A template approach to text analysis: developing and using codebooks. In B. Crabtree, & W. Miller (Eds.), *Doing qualitative research* (pp. 163-177). Newbury Park, CA: Sage.

Dapkus, M. (1985). A thematic analysis of the experience of time. *Journal of Personality and Social Psychology*, 49, 408-419.

Ely, M., Vinz, R., Downing, M., & Anzul, M. (1997). *On writing qualitative research: Living by words*. London: Routledge/Falmer.

Fine, M. (1992). *Disruptive voices: The possibilities for feminist research*. Ann Arbor: University of Michigan Press.

Flick, U. (2014). *An introduction to qualitative research*, 5th ed. London: Sage.

Floersch, J., Longhofer, J.L., Kranke, D. & Townsend, L. (2010). Integrating thematic, grounded theory and narrative analysis. *Qualitative Social Work*, 9(3), 407-425.

Forman, J. & Damschroder, L. (2008). Qualitative content analysis. In L. Jacoby & L.A. Siminoff (Eds.), *Empirical methods for bioethics: A primer* (pp. 39-62). New York: Elsevier.

Fornells-Ambrojo, M., Johns, L., Onwumere, J., Garety, P., Milosh, C. et al. (2017). Experiences of outcome monitoring in service users with psychosis: Findings from an Improving Access to Psychological Therapies for people with Severe Mental Illness (IAPT-SMI) demonstration Site. *British Journal of Clinical Psychology*, 56, 253-272.

Gale, N. K., Heath, G., Cameron, E., Rashid, S. & Redwood, S. (2013). Using the framework method for the analysis of qualitative data in multi-disciplinary health research. *BMC Medical Research Methodology*, 13, <http://www.biomedcentral.com/1471-2288/13/117>

Gallagher, S., Wallace, S., Nathan, Y. & McGrath, D. (2013). 'Soft and fluffy': Medical students' attitudes towards psychology in medical education. *Journal of Health Psychology*, 20(1), 91-101.

Glaser, B. (1992). *Basics of Grounded Theory Analysis: Emergence vs. Forcing*. Mill Valley: Sociology Press.

Gough, B. & Lyons, A. (2016). The future of qualitative research in psychology: Accentuating the positive. *Integrative Psychological and Behavioral Science*, 50, 234-243.

- Guest, G., MacQueen, K., & Namey, E. (2012). *Applied thematic analysis*. Thousand Oaks, CA: Sage.
- Ho, K.H.M., Chiang, V.C.L. & Leung, D. (2017). Hermeneutic phenomenological analysis: The 'possibility' beyond 'actuality' in thematic analysis. *Journal of Advanced Nursing*, 73(7), 1757-1766.
- Jankowski, G., Braun, V. & Clarke, V. (2017). Reflecting on qualitative research, feminist methodologies and feminist psychology: In conversation with Virginia Braun and Victoria Clarke. *Psychology of Women Section Review*, 19(1), 43-55.
- Joffe, H. (2012). Thematic analysis. In D. Harper & A. Thompson (Eds.), *Qualitative methods in mental health and psychotherapy: A guide for students and practitioners* (pp. 209-223). Chichester: Wiley.
- Kaye, L.K., Wall, H.J. & Malone, S.A. (2016). "Turn that frown upside-down": A contextual account of emoticon usage on different virtual platforms. *Computers in Human Behavior*, 60, 463-467.
- Kidder, L., & Fine, M. (1987). Qualitative and quantitative methods: When stories converge. *New directions for program evaluation*, 1987(35), 57-75. 10.1002/ev.1459
- King, N. (2014). Using templates in the thematic analysis of text. In C. Cassell & G. Symon (Eds.), *Essential guide to qualitative methods in organisational research* (pp. 256-270). London: Sage.
- King, N. (2012). Doing template analysis. In G. Symon & C. Cassell (Eds.), *Qualitative organisational research: Core methods and current challenges* (pp. 426-450). London: Sage.
- King, N. & Brooks, J. (2017). Thematic analysis in organisational research. In C. Cassell, A.L. Cunliffe & G. Grandy (Eds.), *The Sage handbook of qualitative business and management research methods* (pp. 219-236). London: Sage.
- Lainson, K., Braun, V. & Clarke, V. (2019). Being both narrative practitioner and academic researcher: A reflection on what thematic analysis has to offer narratively informed research. *International Journal of Narrative Therapy and Community Work*, 4, 86-98.
- Levitt, H.M., Bamberg, M., Creswell, J.W., Frost, D.M., Josselson, R. & Suárez-Orozco, C. (2018). Journal article reporting standards for qualitative primary, qualitative meta-analytic, and mixed methods research in psychology: The APA publications and communications task force report. *American Psychologist*, 73(1), 26-46.
- Levitt, H.M., Motulsky, S.L., Wertz, F.J., Morrow, S.L. & Ponterotto, J.G. (2017). Recommendations for designing and reviewing Qualitative research in psychology: Promoting methodological integrity. *Qualitative Psychology*, 4(1), 2-22.
- Madill, A., & Gough, B. (2008). Qualitative research and its place in psychological science. *Psychological Methods*, 13(3), 254-271.

- McLeod, J. (2001). *Qualitative research in counselling and psychotherapy*. London: Sage.
- Miles, M. B. & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook*, 2nd ed. Thousand Oaks, CA: Sage.
- Nadin, S. & Cassell, C. (2014). Using data matrices. In C. Cassell & G. Symon (Eds.), *Essential guide to qualitative methods in organisational research* (pp. 271-287). London: Sage.
- Newton-John, T.R.O., Ventura, A.D., Mosely, K., Browne, J.L. & Speight, J. (2017). 'Are you sure you're going to have another one of those?': A qualitative analysis of the social control and social support models in type 2 diabetes. *Journal of Health Psychology*, 22(14), 1819-1829.
- Palomäki, J., Laakasuo, M. & Salmela, M. (2013). 'This is just so unfair!': A qualitative analysis of loss-induced emotions and tilting in on-line poker. *International Gambling Studies*, 13(2), 255-270.
- Pidgeon, N., & Henwood, K. (1997a). Using grounded theory in psychological research. In N. Hayes (Ed.), *Doing qualitative analysis in psychology* (pp. 245-273). Hove, UK: Psychology Press.
- Potter, J. (1997). Discourse analysis as a way of analysing naturally occurring talk. In D. Silverman (Ed.), *Qualitative research: Theory, method and practice* (pp. 144-160). London: Sage.
- Ritchie, J. & Lewis, J. (2003). *Qualitative research practice: A guide for social science students and researchers*. London: Sage.
- Ritchie, J. & Spencer, L. (1994). Qualitative data analysis for applied policy research. In A. Bryman & R. G. Burgess (Ed.), *Analysing qualitative data* (pp. 173-194). London: Taylor & Francis.
- Sandlewski, M. & Leeman, J. (2012). Writing usable qualitative health research findings. *Qualitative Health Research*, 22(10), 1404-1413.
- Senders, A., Sando, K., Wahbeh, H., Peterson Hiller, A. & Shinto, L. (2016). Managing psychological stress in the multiple sclerosis medical visit: Patient perspectives and unmet needs, *Journal of Health Psychology*, 21(8), 1676-1687.
- Singer, D. & Hunter, M. (1999). The experience of premature menopause: A thematic discourse analysis. *Journal of Reproductive and Infant Psychology*, 17(1), 63-81.
- Smith, B. (2017). Generalizability in qualitative research: Misunderstandings, opportunities and recommendations for the sport and exercise sciences. *Qualitative Research in Sport, Exercise and Health*, 10(1), 137-149.
- Smith, J. & Firth, J. (2011). Qualitative data approaches: The framework approach. *Nurse Researcher*, 18(2), 52-62.

- Smith, J. A., Flowers, P., & Larkin, M. (2009). *Interpretative Phenomenological Analysis: Theory, Method and Research*. London: Sage.
- Sparkes, A.C. & Smith, B. (2009). Judging the quality of qualitative inquiry: Criteriology and relativism in action. *Psychology of Sport and Exercise*, 10, 491-497.
- St. Pierre, E. A. (2019). Post Qualitative Inquiry, the Refusal of Method, and the Risk of the New. *Qualitative Inquiry*, online first.
- St. Pierre, E. A., & Jackson, A. Y. (2014). Qualitative Data Analysis After Coding. *Qualitative Inquiry*, 20(6), 715-719.
- Stenner, P. (1993). Discoursing jealousy. In E. Burman & I. Parker (Eds.), *Discourse analytic research: Repertoires and readings of texts in action* (pp. 94-132). London: Routledge.
- Taylor, G. W., & Ussher, J. M. (2001). Making sense of S&M: A discourse analytic account. *Sexualities*, 4(3), 293-314.
- Terry, G. & Braun, V. (2011). "It's kind of me taking responsibility for these things": Men, vasectomy and contraceptive economies. *Feminism & Psychology*, 21(4), 477-495.
- Terry, G., Hayfield, N., Clarke, V. & Braun, V. (2017). Thematic analysis. In Willig, C. & Stainton-Rogers (Eds.), *The Sage handbook of qualitative research in psychology*, 2nd edition (pp. 17-37). London: Sage.
- Tischner, I. (2019). Tomorrow is the start of the rest of their life — so who cares about health? Exploring constructions of weight-loss motivations and health using story completion. *Qualitative Research in Psychology*, 16(1), 54-73.
- Vaismoradi, M., Turunen, H. & Bondas, T. (2013). Content analysis and thematic analysis: Implications for conducting a qualitative descriptive study. *Nursing & Health Sciences*, 15, 398-405.
- Wetherell, M., & Edley, N. (2009). *Masculinity manoeuvres: Critical discourse psychology and the analysis of identity strategies* (pp. 201-214). Basingstoke: Palgrave Macmillan.
- Willcox, R., Moller, N. & Clarke, V. (2019). Exploring attachment incoherence in bereaved families' therapy narratives: An attachment theory-informed thematic analysis. *The Family Journal: Counseling and Therapy for Couples and Families*, 27(3), 339-347.
- Willig, C. (2013). *Introducing qualitative research in psychology*, 3rd ed. Berkshire: Open University Press.
- Yardley, L. (2015). Demonstrating validity in qualitative psychology. In J.A. Smith (Ed.), *Qualitative psychology: A practical guide to research methods*, 2nd ed. (pp. 257-272). London: Sage.

Table 1: A tool for evaluating thematic analysis (TA) manuscripts for publication: Twenty questions to guide assessment of TA research quality

These questions are designed to be used either independently, or alongside our methodological writing on TA, and especially the current paper, if further clarification is needed.

Adequate choice and explanation of methods and methodology

1. Do the authors explain why they are using thematic analysis (TA), even if only briefly?
2. Do the authors clearly specify and justify which *type* of TA they are using?
3. Is the use and justification of the specific type of TA consistent with the research questions or aims?
4. Is there a good 'fit' between the theoretical and conceptual underpinnings of the research and the specific type of TA (i.e. is there conceptual coherence)?
5. Is there a good 'fit' between the methods of data collection and the specific type of TA?
6. Is the specified type of TA consistently enacted throughout the paper?
7. Is there evidence of problematic assumptions about, and practices around, TA?
These commonly include:
 - Treating TA as one, homogenous, entity, with one set of – widely agreed on – procedures.
 - Combining philosophically and procedurally incompatible approaches to TA without any acknowledgement or explanation.
 - Confusing summaries of data topics with thematic patterns of shared meaning, underpinned by a core concept.
 - Assuming grounded theory concepts and procedures (e.g. saturation, constant comparative analysis, line-by-line coding) apply to TA without any explanation or justification.
 - Assuming TA is essentialist or realist, or atheoretical.
 - Assuming TA is only a data reduction or descriptive approach and therefore must be supplemented with other methods and procedures to achieve other ends.
8. Are any supplementary procedures or methods justified, and necessary, or could the same results have been achieved simply by using TA more effectively?
9. Are the theoretical underpinnings of the use of TA clearly specified (e.g. ontological, epistemological assumptions, guiding theoretical framework(s)), even when using TA inductively (inductive TA does not equate to analysis in a theoretical vacuum)?

10. Do the researchers strive to 'own their perspectives' (even if only very briefly), their personal and social standpoint and positioning? (This is especially important when the researchers are engaged in social justice-oriented research and when representing the 'voices' of marginal and vulnerable groups, and groups to which the researcher does not belong.)
11. Are the analytic procedures used clearly outlined, and described in terms of what the authors actually did, rather than generic procedures?
12. Is there evidence of conceptual and procedural confusion? For example, reflexive TA (Braun & Clarke, 2006) is the claimed approach but different procedures are outlined such as the use of a codebook or coding frame, multiple independent coders and consensus coding, inter-rater reliability measures, and/or themes are conceptualised as analytic inputs rather than outputs and therefore the analysis progresses from theme identification to coding (rather than coding to theme development).
13. Do the authors demonstrate full and coherent understanding of their claimed approach to TA?

A well-developed and justified analysis

14. Is it clear what and where the themes are in the report? Would the manuscript benefit from some kind of overview of the analysis: listing of themes, narrative overview, table of themes, thematic map?
15. Are reported themes topic summaries, rather than 'fully realised themes' – patterns of shared meaning underpinned by a central organising concept?
 - Have the data collection questions been used as themes?
 - If so, are topic summaries appropriate to the purpose of the research?
 - If the authors are using reflexive TA, is this modification in the conceptualisation of themes explained and justified?
 - Would the manuscript benefit from further analysis being undertaken, with the reporting of fully realised themes?
 - Or, if the authors are claiming to use reflexive TA, would the manuscript benefit from claiming to use a different type of TA (e.g. coding reliability or codebook)?
16. Is a non-thematic contextualising information presented as a theme? (e.g. the first theme is a topic summary providing contextualising information, but the rest of the themes reported are fully realised themes). If so, would the manuscript benefit from this being presented as non-thematic contextualising information?
17. In applied research, do the reported themes have the potential to give rise to actionable outcomes?

18. Are there conceptual clashes and confusion in the paper? (e.g. claiming a social constructionist approach while also expressing concern for positivist notions of coding reliability, or claiming a constructionist approach while treating participants' language as a transparent reflection of their experiences and behaviours)

19. Is there evidence of weak or unconvincing analysis such as:

- Too many or too few themes?
- Too many theme levels?
- Confusion between codes and themes?
- Mismatch between data extracts and analytic claims?
- Too few or too many data extracts?
- Overlap between themes?

20. Do authors make problematic statements about the lack of generalisability of their results, and or implicitly conceptualise generalisability as statistical probabilistic generalisability (see Smith, 2017)?